

Datasets and Interfaces for Benchmarking Heterogeneous Graph Neural Networks

Yijian Liu
Beijing Key Lab of Intelligent
Telecommunication Software and
Multimedia
Beijing University of Posts and
Telecommunications
Beijing, China
liuyijian@bupt.edu.cn

Luhao Zhang
Tao Li
zhangluhao@meituan.com
litaot19@meituan.com
Meituan
Beijing, China

Hongyi Zhang
Cheng Yang
zhanghongyi@bupt.edu.cn
yangcheng@bupt.edu.cn
Beijing University of Posts and
Telecommunications
Beijing, China

Jinyu Yang
Tianyu Zhao
Juan Yang*
jinyu.yang@bupt.edu.cn
tyzhao@bupt.edu.cn
yangjuan@bupt.edu.cn
Beijing University of Posts and
Telecommunications
Beijing, China

Ao Li
Yugang Ji
byr.li.ao@gmail.com
jiyugang9305@gmail.com
Orange Shield Technology
Hangzhou, China

Hai Huang
Chuan Shi*
hhuang@bupt.edu.cn
shichuan@bupt.edu.cn
Beijing University of Posts and
Telecommunications
Beijing, China

ABSTRACT

In recent years, Heterogeneous Graph Neural Networks (HGNNs) have gained increasing attention due to their excellent performance in applications. However, the lack of high-quality benchmarks in new fields has become a critical limitation for developing and applying HGNNs. To accommodate the urgent need for emerging fields and the advancement of HGNNs, we present two large-scale, real-world, and challenging heterogeneous graph datasets from real scenarios: risk commodity detection and takeout recommendation. Meanwhile, we establish standard benchmark interfaces that provide over 40 heterogeneous graph datasets. We provide initial data split, unified evaluation metrics, and baseline results for future work, making it fair and handy to explore state-of-the-art HGNNs. Our interfaces also offer a comprehensive toolkit to research the characteristics of graph datasets. The above new datasets are publicly available on <https://zenodo.org/communities/hgd>, and the interface codes are available at <https://github.com/BUPT-GAMMA/hgbi>.

CCS CONCEPTS

• **Information systems** → *Information systems applications; Data mining*; • **Computing methodologies** → **Artificial intelligence**;

*The corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CIKM '23, October 21–25, 2023, Birmingham, United Kingdom
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0124-5/23/10...\$15.00
<https://doi.org/10.1145/3583780.3615117>

KEYWORDS

Heterogeneous Graph Neural Networks, Graph, Benchmark, Risk Commodity Detection, Takeout Recommendation

ACM Reference Format:

Yijian Liu, Hongyi Zhang, Cheng Yang, Ao Li, Yugang Ji, Luhao Zhang, Tao Li, Jinyu Yang, Tianyu Zhao, Juan Yang, Hai Huang, and Chuan Shi. 2023. Datasets and Interfaces for Benchmarking Heterogeneous Graph Neural Networks. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM '23)*, October 21–25, 2023, Birmingham, United Kingdom. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3583780.3615117>

1 INTRODUCTION

Traditional deep learning primarily focuses on learning the representation of Euclidean data (e.g., text and images). However, data in the real world is often more complex and exists in irregular formats. A heterogeneous graph is one such irregular data that comprises multi-type nodes and multi-type edges such as papers, authors, subjects, and terms in a publication network [20]. Heterogeneous Graph Neural Networks (HGNNs) is an efficient technique to utilize complex graph structures and capture latent semantic information for heterogeneous graphs. Over the past few years, HGNNs have achieved notable success and demonstrated evidence of their efficiency and effectiveness in a variety of applications [5, 12, 13, 15, 16, 24], including user behavior analysis, drug molecule design, product recommendation, and more.

Benchmarks play a crucial role in exploring state-of-the-art HGNNs in heterogeneous graphs, and they are driving applications of HGNNs in downstream tasks. However, most available graph datasets are small-scale or homogeneous, e.g., Cora, Citeseer, and Reddit [6, 14, 30], which suffer from poor quality such as duplicate

nodes and data leakage. Additionally, traditional deep learning algorithms have already achieved high performance on these datasets, which serve no purpose for exploring state-of-the-art methods of HGNNs. Large-scale and heterogeneous graph datasets provided by public benchmarks, such as OGB [10], HNE [26] and HGB [17], mainly concentrate on the mundane research scenarios or common life scenarios, *e.g.*, paper citation network, protein/drug association network, and general product recommendation. As new fields rapidly emerge, the existing graph datasets fall short of containing their new and unique characteristics, posing new challenges for HGNNs in these untrodden fields. Therefore, it is urgent and necessary to establish new datasets to reflect real-world scenarios. Moreover, previous graph datasets have also been hampered by several shortcomings, including the lack of available codes, loading interfaces or unified metrics, and explicit separation of datasets and tasks, all of which lead to errors, unfairness, and complexity during model evaluation. To address the above limitations of previous graph datasets, our main contributions are the following:

Large-scale graph datasets. We present two large-scale heterogeneous graph datasets in emerging and thriving fields: risk commodity detection and takedown recommendation. We provide data split, unified metrics, and results of baselines. In addition, we hold a corresponding competition to facilitate state-of-the-art HGNNs.

Benchmark interfaces. We reproduce previous work and establish benchmark interfaces that provide over 40 heterogeneous graph datasets sourced from other fields. Our interfaces offer a standardized way to load, process, and customize graph datasets. Towards advanced graph tasks such as node classification and link prediction, we provide unified metrics and benchmark HGNNs. Besides that, our interfaces also offer a comprehensive toolkit to analyze the characteristics of graph datasets.

2 DATASETS

2.1 Risk Commodity Detection Dataset

2.1.1 Practical Relevance and Challenge. The existence of risk commodities is gradually endangering e-commerce platforms. To tackle this issue, we establish a risk commodity detection dataset (RCDD) based on a real risk detection scenario from Alibaba's e-commerce platform. In this detection scenario, risk commodities always deliberately disguise risk information, leading to a fierce confrontation between them and risk control systems. The main challenges we face include malicious users who forge "innocent" relationships by forging devices, addresses, or other methods. Moreover, the distribution of black and white samples is severely imbalanced, and the graph is large-scale and heterogeneous.

2.1.2 Graph. This graph dataset is manually labeled by ourselves to ensure reliable ground truth. Based on the breadth-first search, we start from risk commodity nodes and spread outward to sample other nodes. The schema diagram of the graph is illustrated in Figure 1, which is large-scale with 157,814,864 edges and 13,806,619 nodes. Due to the sensitivity of this scenario, for confidentiality and security, except commodity, we cannot disclose the names of the other node types (*e.g.*, buyer and seller) and edge types (*e.g.*,

buy and sell) which are represented by single letters. The 256-dimension feature of the commodity is concatenated by the image and text features extracted from pre-trained models, BERT [2] and BYOL [4], respectively. Features of other node types are obtained by averaging the features of their neighbors. Due to the excessive number of white samples in the dataset, we downsample them to decrease the amount of them. In data split, the validation set is split from the training set, and the test set is obtained over time.

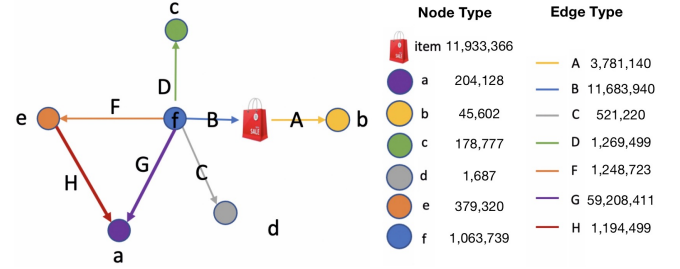


Figure 1: A schema diagram of RCDD.

2.1.3 Metric and Baseline. The graph task is node classification: detect risk commodities. In Table 1, we benchmark a broad range of 7 powerful HGNN models with 3 layers, and the evaluation metrics are Macro/Micro-F1 and AP (average precision). We are surprised to find that general HGNNs (*e.g.*, RGCN and Rsage) outperform the later specifically designed HGNNs (*e.g.*, ieHGCN and SimpleHGN).

Table 1: Baseline results in RCDD.

Model	#Params	Macro/Micro-F1 (%)	AP (%)
RGCN [19]	231,746	90.46/98.02	87.72
Rsage [6]	524,636	89.99/98.10	87.21
ieHGCN [29]	588,208	88.95/97.60	87.06
SimpleHGN [17]	1,638,086	88.70/97.88	83.97
RGAT [22]	835,772	88.54/97.65	83.79
HAN [25]	1,034,370	84.67/97.01	77.71
HetSANN [18]	101,650	82.12/96.84	73.22

2.1.4 Competition. We hold a risk commodities detection competition that has attracted more than 2000 teams globally. Despite enormous challenges in this new scenario, the top three teams still achieve remarkable scores: 94.64%, 94.54%, and 94.22% on AP, all of which outperform our baselines significantly. Upon reviewing and evaluating the technical reports from these excellent teams, we discover some innovative and practical attempts, such as artificially introducing noise/disturbance to attain adversarial training and performing data augmentation for the neighbor nodes of the commodity nodes, which solves the issue of imbalanced distribution. Overall, it is crucial to maintain the health and stability of the e-commerce industry, and RCDD is a challenging graph dataset that can realistically reflect the scenario.

2.2 Takeout Recommendation Dataset

2.2.1 Practical Relevance and Challenge. The takeout industry has brought huge convenience and impacted people’s lives greatly. However, the massive amount of takeout restaurants and food information has been bothering the users since their attention is limited. Our takeout recommendation dataset (**TRD**) [27, 28, 34] is sourced from the Meituan Takeout app and presents new challenges compared to traditional product recommendation datasets. These challenges include multi-dimensional attributes, short decision time, cyclical user interests, and diverse food. The most challenging aspect of this dataset is that it contains spatial and temporal information.

2.2.2 Graph. We collect orders from 11 commercial districts in Beijing from March 1st to March 28th, 2021. The first three weeks of orders are as training, while the last week is used for testing. On data cleaning, we remove dirty data and exclude irrelevant information whose number of orders is less than two and those who continuously order the same food at the same takeout restaurant. Besides that, we encrypt sensitive information to protect user privacy. The graph constructed by ourselves is illustrated in Figure 2, where *poi* is the takeout restaurant, and *spu* is food, and this graph is huge with 18,931,400 edges and 408,849 nodes. Unlike other simple graphs, our graph integrates a vast amount of meta information on nodes and edges, which is a novel characteristic. We record essential data such as consumption information for each user, as well as price, taste, and ingredients for each food. For each takeout restaurant, we record location information and scores of food, delivery service, and overall service. For each order, we collect the timestamp and the user’s delivery address.

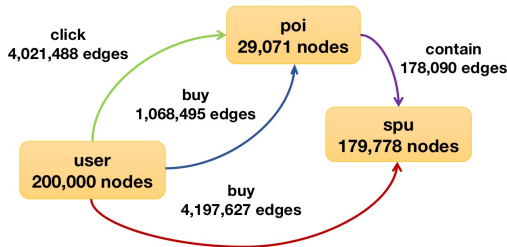


Figure 2: A schema diagram of TRD.

2.2.3 Metric and Baseline. The graph task is link prediction, which predicts whether there is an edge between food and the user. The evaluation metric is AUC-ROC (area under the roc curve), and the results of baselines are shown in Table 2. We can see that HGNNs have demonstrated powerful potential in this field, making them possible to apply in the real world.

3 INTERFACES FOR BENCHMARKING

We offer standardized interfaces to load, process, and store graphs. Users no longer need to download any raw file, preprocess, and then manually convert it to graph format. Over 40 heterogeneous graph datasets are readily available to users, which support not only HGNNs but also a brand range of other graph learning algorithms, making data acquisition and manipulation a seamless experience.

Table 2: Baseline results in TRD.

Model	#Params	AUC-ROC (%)
RGCN	26,264,832	92.69
RGAT	26,582,808	91.53
CompGCN [21]	13,091,968	89.88

3.1 Load

We adhere to the unification of graph dataset and task [7], ensuring that each graph dataset corresponds to a specified task, thereby simplifying the application of downstream tasks. As the code block is shown below, users can easily load the dataset by selecting the dataset name and task name. The dataset class contains essential attributes such as meta path, category, save path, and graph format in DGL [23], all of which ensure the integrity of the dataset. Table 3 and Table 4 summarize some of the graph datasets through our interfaces for different graph tasks.

```

import hgbi
ds_node = hgbi.build_dataset( #risk commodity detection
    name = 'RCDD', task = 'node_classification')
ds_link = hgbi.build_dataset( #takeout recommendation
    name = 'TRD', task = 'link_prediction')

```

3.2 Customization

Our interfaces allow users to load their graph files by our built-in class. Further customization is also available through our task adaptor, including specifying labeled node type, data split, negative sampling, adding reverse edges, etc.

```

my_ds = hgbi.MyDataset(path="./graph.bin")
my_ds_node = hgbi.AsNodeClassificationDataset(my_ds,
    labeled_nodes_split_ratio=[0.5,0.3,0.3],
    target_ntype="node-1",
    label_feat_name='label')
my_ds_link = hgbi.AsLinkPredictionDataset(my_ds,
    target_link=['edge-1'], split_ratio=[0.5,0.3,0.3],
    target_link_r=['rev_edge-1'], neg_ratio=3,
    neg_sampler='global')

```

3.3 Evaluation

To evaluate HGNNs, we provide unified metrics across all graph datasets, Macro/Micro-F1 for node classification, and AUC-ROC for link prediction. The reproduced results of baselines are also shown in Table 3 and Table 4.

3.4 Toolkit

In this section, we explore the critical characteristics of graph datasets and present a powerful and innovative toolkit to analyze graph datasets. This toolkit can give a novel overview of graph datasets, especially on heterogeneous graphs.

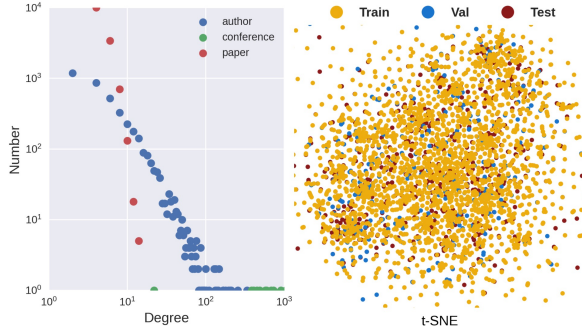
3.4.1 Degree distribution of multi-type nodes. The analysis of the graph’s degree distribution can provide insights into its topology and the relative significance of nodes. In the left panel of Figure 3, our toolkit shows a clear visualization example for degree distribution.

Table 3: Statistics of graph datasets and reproduced results of Macro/Micro-F1 (%) on node classification.

Dataset	Ntype	Node	Etype	Edge	Avg Attr	Label	Model	Reproduced
acm4NARS	3	21,488	4	34,864	720	3	NARS [31]	91.35/91.44
acm4HetGNN	3	49,708	5	202,067	387	4	HetGNN [33]	97.01/97.05
imdb4MAGNN	3	11,616	4	34,212	3,468	3	MAGNN [3]	62.85/62.78
dblp4GTN	3	18,405	4	67,946	334	4	fastGTN [32]	90.39/91.39
yelp4HeGAN	5	3,913	8	77,360	64	3	HeGAN [8]	71.51/79.16
HGBn-DBLP	4	26,128	6	239,566	1,538	4	SimpleHGN	86.31/87.24
ohgbn-Freebase	8	12,164,755	36	62,982,566	N/A	8	RGCN	53.07/69.33
ohgbn-yelp2	4	82,465	4	30,542,675	N/A	16	RGCN	5.04/40.44
RCDD	7	13,806,619	7	157,814,864	256	2	RGCN	90.46/98.02

Table 4: Statistics of graph datasets and reproduced results of AUC-ROC (%) on link prediction.

Dataset	Ntype	Node	Etype	Edge	Avg Attr	Label	Model	Reproduced
amazon4SLICE	1	10,099	2	170,783	1,156	2	RGCN	74.60
HGBI-DBLP	4	26,128	6	239,566	1,538	1	HDE [11]	98.36
HGBI-IMDB	4	21,420	6	86,642	3,390	1	HDE	91.51
HGBI-amazon	1	10,099	2	148,659	1,156	2	GATNE-T [1]	80.83
HGBI-LastFM	3	20,612	6	283,042	N/A	1	RGCN	79.88
HGBI-PubMed	4	63,109	20	489,972	200	1	RGCN	89.30
DoubanMovie	6	37,595	12	3,429,852	N/A	1	RGCN	91.55
TRD	3	408,849	4	18,931,400	N/A	1	RGCN	92.69

**Figure 3: Degree distribution (left) and t-SNE visualization (right) in dblp4GTN.**

3.4.2 Visualization of data split. Different data splits may result in a vast generalization gap in experiments [10]. We provide t-SNE visualization, shown in the right panel of Figure 3, which can reflect data split and bring more interpretability to experiments.

3.4.3 Connection strength and heterophily. Given a meta-path, nodes are regarded as neighbors if they are along it. Connection strength is a threshold that indicates the number of edges along one meta-path must be larger than the threshold. We extend the concept of homophily [9] to heterophily for heterogeneous graphs, whereby heterophily is the ratio of 2 nodes on the meta-path having different target labels. In Table 5, where *A* is the author, *P* is the paper, and *C* is the conference, this example analyzes the connection strength and heterophily for meta paths in dblp4GTN. Heterophily is a vital property to reflect the characteristics of heterogeneous graphs, which can improve significant performance for HGNNs.

Table 5: Connection strength and Heterophily in dblp4GTN.

Meta-path	Connctcion Strength (%)	Heterophily (%)	Edge Ratio (%)
A-P-A	2	7.93	50.53
	4	6.27	80.17
	8	3.98	93.22
A-P-C-P-A	2	30.49	31.36
	4	27.41	62.38
	8	25.36	82.02

4 CONCLUSION AND FUTURE WORK

This paper presents two heterogeneous graph datasets with novel challenges and characteristics in new emerging fields: risk commodity detection and takeout recommendation. These two graphs are indicative of real-world applications and can reflect the reality of scenarios. In addition, We establish benchmark interfaces that contain more than 40 heterogeneous graph datasets, and we offer unified metrics to benchmark HGNNs. We explore the characteristics of the graph datasets in multiple aspects through our interfaces. In the future, we will continue to maintain the community for graph benchmarks and support researchers in sharing graph datasets.

ACKNOWLEDGMENTS

This work is supported in part by the National Natural Science Foundation of China (No. 62192784, U1936104, U20B2045, 62172052, 62002029).

REFERENCES

- [1] Yukuo Cen, Xu Zou, Jianwei Zhang, Hongxia Yang, Jingren Zhou, and Jie Tang. 2019. Representation learning for attributed multiplex heterogeneous network. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 1358–1368.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).
- [3] Xinyu Fu, Jiani Zhang, Ziqiao Meng, and Irwin King. 2020. Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding. In *Proceedings of The Web Conference 2020*. 2331–2341.
- [4] Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, et al. 2020. Bootstrap your own latent—a new approach to self-supervised learning. *Advances in neural information processing systems* 33 (2020), 21271–21284.
- [5] Shuyun Gu, Xiao Wang, Chuan Shi, and Ding Xiao. 2022. Self-supervised Graph Neural Networks for Multi-behavior Recommendation. In *International Joint Conference on Artificial Intelligence (IJCAI)*.
- [6] Will Hamilton, Zhitaoying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems* 30 (2017).
- [7] Hui Han, Tianyu Zhao, Cheng Yang, Hongyi Zhang, Yaoqi Liu, Xiao Wang, and Chuan Shi. 2022. Openhgnn: an open source toolkit for heterogeneous graph neural network. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 3993–3997.
- [8] Binbin Hu, Yuan Fang, and Chuan Shi. 2019. Adversarial learning on heterogeneous information networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 120–129.
- [9] Weihua Hu, Matthias Fey, Hongyu Ren, Maho Nakata, Yuxiao Dong, and Jure Leskovec. 2021. Ogb-lsc: A large-scale challenge for machine learning on graphs. *arXiv preprint arXiv:2103.09430* (2021).
- [10] Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, and Jure Leskovec. 2020. Open graph benchmark: Datasets for machine learning on graphs. *Advances in neural information processing systems* 33 (2020), 22118–22133.
- [11] Houye Ji, Cheng Yang, Chuan Shi, and Pan Li. 2021. Heterogeneous graph neural network with distance encoding. In *2021 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1138–1143.
- [12] Yugang Ji, Guanyi Chu, Xiao Wang, Chuan Shi, Jianan Zhao, and Junping Du. 2022. Prohibited Item Detection via Risk Graph Structure Learning. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, Frédérique Laforest, Raphaël Troncy, Elena Simperl, Deepak Agarwal, Aristides Gionis, Ivan Herman, and Lionel Médini (Eds.). ACM, 1434–1443. <https://doi.org/10.1145/3485447.3512190>
- [13] Yugang Ji, Chuan Shi, and Xiao Wang. 2021. Prohibited Item Detection on Heterogeneous Risk Graphs. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1 - 5, 2021*, Gianluca Demartini, Guido Zuccon, J. Shane Culpepper, Zi Huang, and Hanghang Tong (Eds.). ACM, 3867–3877. <https://doi.org/10.1145/3459637.3481945>
- [14] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
- [15] Ao Li, Yugang Ji, Guanyi Chu, Xiao Wang, Dong Li, and Chuan Shi. 2023. Clustering-Based Supervised Contrastive Learning for Identifying Risk Items on Heterogeneous Graph. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.
- [16] Jiawei Liu, Chuan Shi, Cheng Yang, Zhiyuan Lu, and S Yu Philip. 2022. A survey on heterogeneous information network based recommender systems: Concepts, methods, applications and resources. *AI Open* 3 (2022), 40–57.
- [17] Qingsong Lv, Ming Ding, Qiang Liu, Yuxiang Chen, Wenzheng Feng, Siming He, Chang Zhou, Jianguo Jiang, Yuxiao Dong, and Jie Tang. 2021. Are we really making much progress? revisiting, benchmarking and refining heterogeneous graph neural networks. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 1150–1160.
- [18] M Marvi-Mashhadi, CS Lopes, and J Llorca. 2020. High fidelity simulation of the mechanical behavior of closed-cell polyurethane foams. *Journal of the Mechanics and Physics of Solids* 135 (2020), 103814.
- [19] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings 15*. Springer, 593–607.
- [20] Chuan Shi, Yitong Li, Jiawei Zhang, Yizhou Sun, and Philip S. Yu. 2017. A survey of heterogeneous information network analysis. *IEEE Transactions on Knowledge and Data Engineering* 29, 1 (2017), 17–37. <https://doi.org/10.1109/TKDE.2016.2598561>
- [21] Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, and Partha Talukdar. 2019. Composition-based multi-relational graph convolutional networks. *arXiv preprint arXiv:1911.03082* (2019).
- [22] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).
- [23] Minjie Wang, Da Zheng, Zihao Ye, Quan Gan, Mufei Li, Xiang Song, Jinjing Zhou, Chao Ma, Lingfan Yu, Yu Gai, et al. 2019. Deep graph library: A graph-centric, highly-performant package for graph neural networks. *arXiv preprint arXiv:1909.01315* (2019).
- [24] Xiao Wang, Deyu Bo, Chuan Shi, Shaohua Fan, Yanfang Ye, and S Yu Philip. 2022. A survey on heterogeneous graph embedding: methods, techniques, applications and sources. *IEEE Transactions on Big Data* (2022).
- [25] Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, and Philip S Yu. 2019. Heterogeneous graph attention network. In *The world wide web conference*. 2022–2032.
- [26] Carl Yang, Yuxin Xiao, Yu Zhang, Yizhou Sun, and Jiawei Han. 2020. Heterogeneous network representation learning: A unified framework with survey and benchmark. *IEEE Transactions on Knowledge and Data Engineering* 34, 10 (2020), 4854–4873.
- [27] Tianchi Yang, Cheng Yang, Luhao Zhang, Chuan Shi, Maodi Hu, Huaijun Liu, Tao Li, and Dong Wang. 2022. Co-clustering interactions via attentive hypergraph neural network. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 859–869.
- [28] Tianchi Yang, Luhao Zhang, Chuan Shi, Cheng Yang, Siyong Xu, Ruiyu Fang, Maodi Hu, Huaijun Liu, Tao Li, and Dong Wang. 2022. Gated Hypergraph Neural Network for Scene-Aware Recommendation. In *International Conference on Database Systems for Advanced Applications*. Springer, 199–215.
- [29] Yaming Yang, Ziyu Guan, Jianxin Li, Wei Zhao, Jiangtao Cui, and Quan Wang. 2021. Interpretable and efficient heterogeneous graph convolutional network. *IEEE Transactions on Knowledge and Data Engineering* (2021).
- [30] Zhilin Yang, William Cohen, and Ruslan Salakhudinov. 2016. Revisiting semi-supervised learning with graph embeddings. In *International conference on machine learning*. PMLR, 40–48.
- [31] Lingfan Yu, Jiajun Shen, Jinyang Li, and Adam Lerer. 2020. Scalable graph neural networks for heterogeneous graphs. *arXiv preprint arXiv:2011.09679* (2020).
- [32] Seongjun Yun, Minbyul Jeong, Sungdong Yoo, Seunghun Lee, S Yi Sean, Raehyun Kim, Jaewoo Kang, and Hyunwoo J Kim. 2022. Graph Transformer Networks: Learning meta-path graphs to improve GNNs. *Neural Networks* 153 (2022), 104–119.
- [33] Chuxu Zhang, Dongjin Song, Chao Huang, Ananthram Swami, and Nitesh V Chawla. 2019. Heterogeneous graph neural network. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 793–803.
- [34] Luhao Zhang, Ruiyu Fang, Tianchi Yang, Maodi Hu, Tao Li, Chuan Shi, and Dong Wang. 2022. A Joint Framework for Explainable Recommendation with Knowledge Reasoning and Graph Representation. In *International Conference on Database Systems for Advanced Applications*. Springer, 351–363.